Multiplicity distributions in pp collisions from STAR experiment

Jan Kapitán
(for the STAR Collaboration)

Nuclear Physics Institute ASCR
Czech Republic

Outline and motivation

- get the corrected charged multiplicity distribution (using an unfolding method, based on Bayes' theorem)

- reach much higher multiplicities than the previous measurement (UA5 at $S\bar{p}pS$)

- predict multiplicity distributions for LHC energies

Multiplicity distributions

- Event charged multiplicity:
  - N_{ch} (or simply N) – number of charged tracks coming from event primary vertex (primary tracks)
  - weak decay and gamma conversion products have to be rejected
- UA5 experiment (1985): multiplicity distributions follow the Negative Binomial Distribution
- UA5 measured Non Singly Diffractive (NSD) events

STAR experiment

(Solenoidal Tracker At Rhic, Brookhaven National Lab, NY)
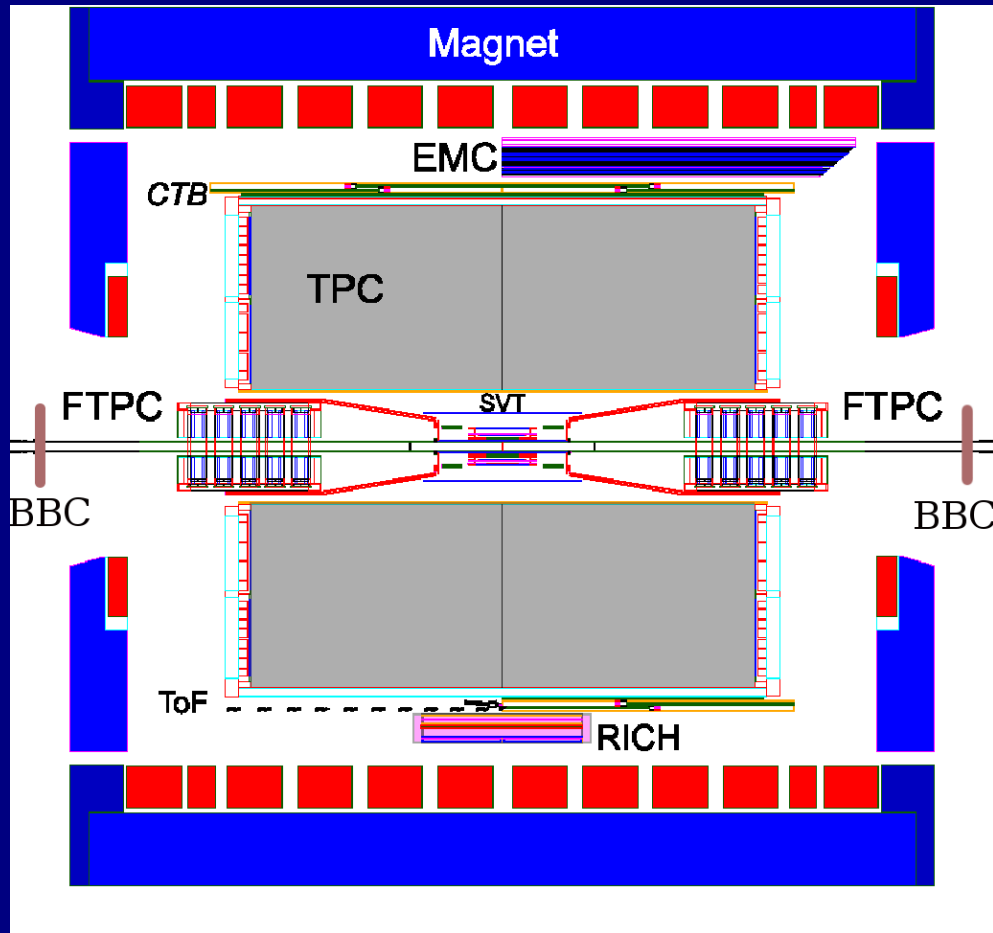
- Time Projection Chamber:
  - main tracker ($p_T > 0.1$ GeV/c and $|\eta| < 1.8$)
  - measures dE/dx --> PID
- BBC trigger detector – minimum bias trigger, coincidence, NSD
- CTB detector – track matching: pile-up rejection

Data analysis

data: minimum bias pp collisions at $\sqrt{s} = 200$ GeV

↓

event cuts: vertex $|z_{vertex}| < 25$ cm – center of the TPC

track selection: $|\eta| < 0.5$, $p_T > 0.15$ GeV/c, track quality cuts, DCA to primary vertex < 1 cm

↓

vertex finding efficiency correction (low multiplicity events !)
correction for tracking efficiency + contamination (weak decays and gamma conversions)

↓

estimate systematic uncertainties from the corrections

Vertex finding efficiency

- depends on the number of tracks reconstructed in the TPC
- each event weighed by the inverse probability
- CTB matching: pile-up rejection
- obtained from the data
- decreases $<N_{ch}>$ from 2.11 to 1.95
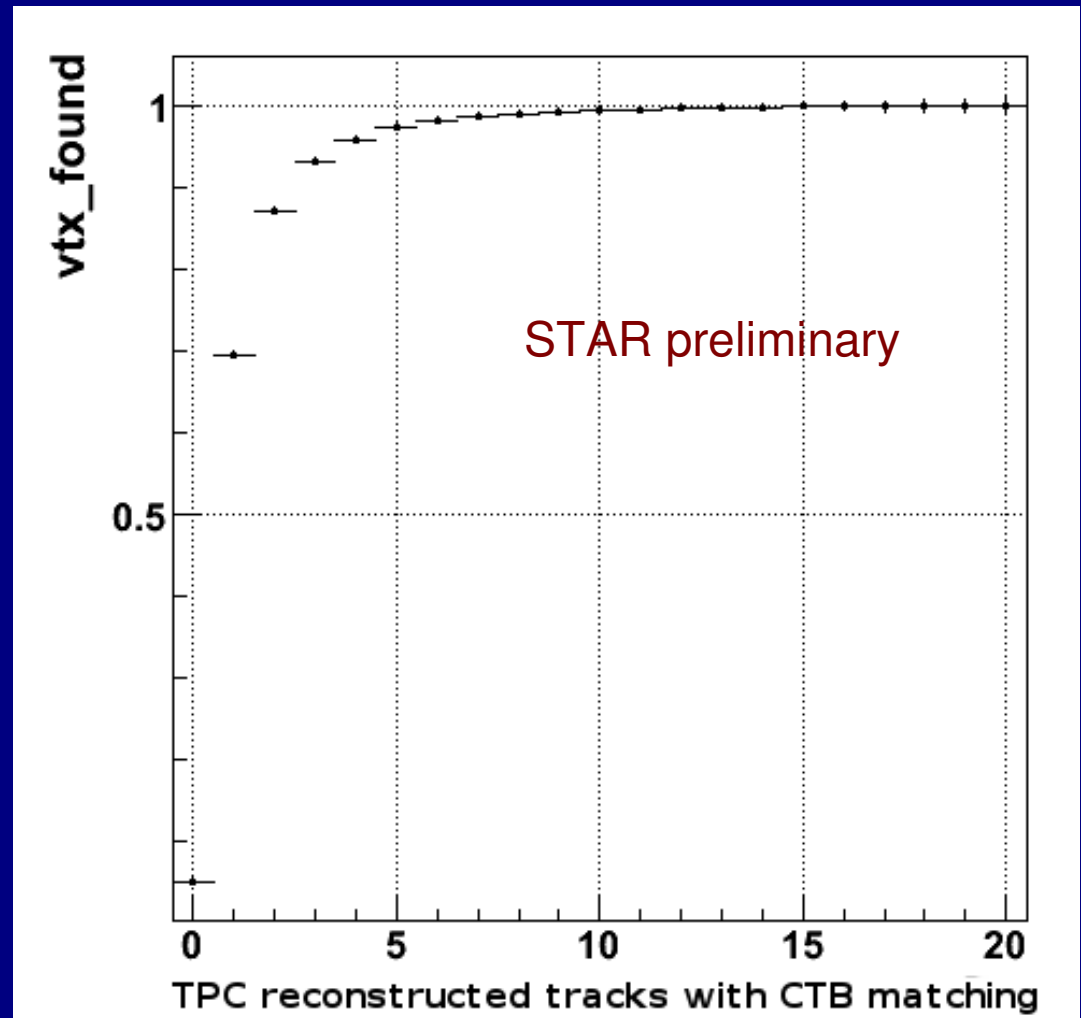
Tracking efficiency & contamination

- Hijing (pp) events, GEANT
- embedded into realistic background
- reconstruction & association
- the same cuts as in data analysis

- H: multiplicity from the simulation (before GEANT)
- G: reconstructed multiplicity ("after GEANT primaries")
- $P(G|H)$: smearing matrix

STAR preliminary
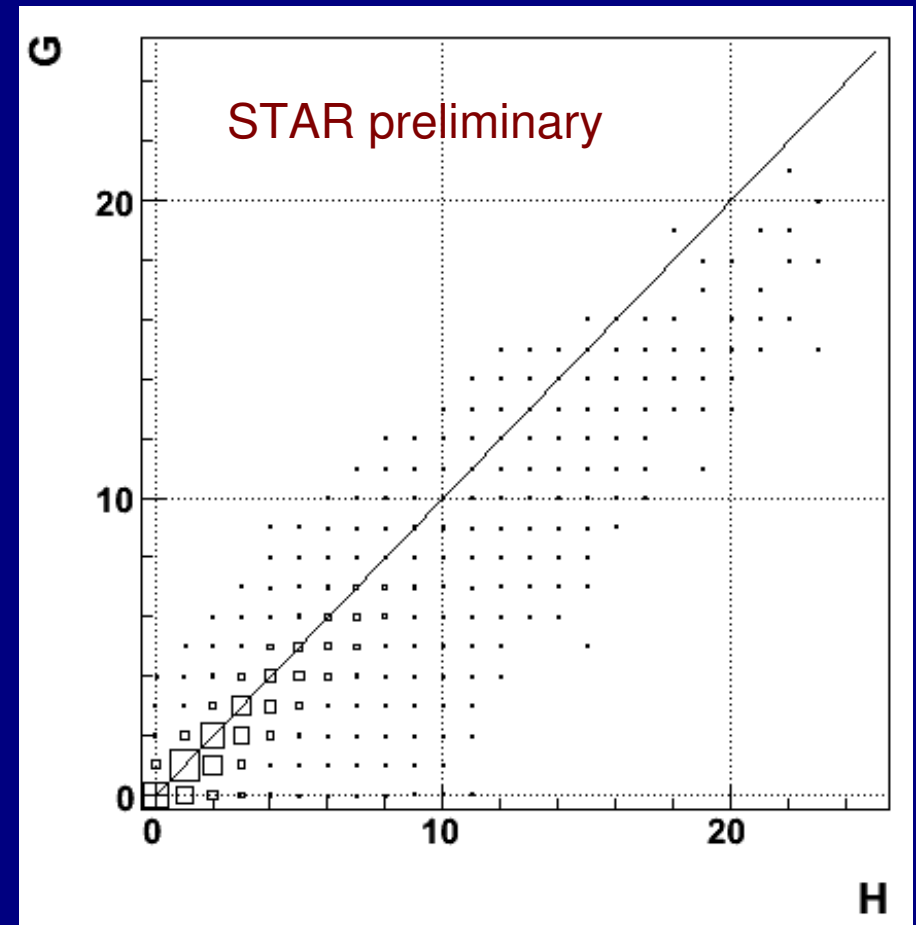
average "efficiency" 82 %: will increase $<N_{ch}>$ from 1.95 to 2.38

Contamination: for DCA < 1 cm, 8% of "after GEANT primaries" come from weak decays / gamma conversions
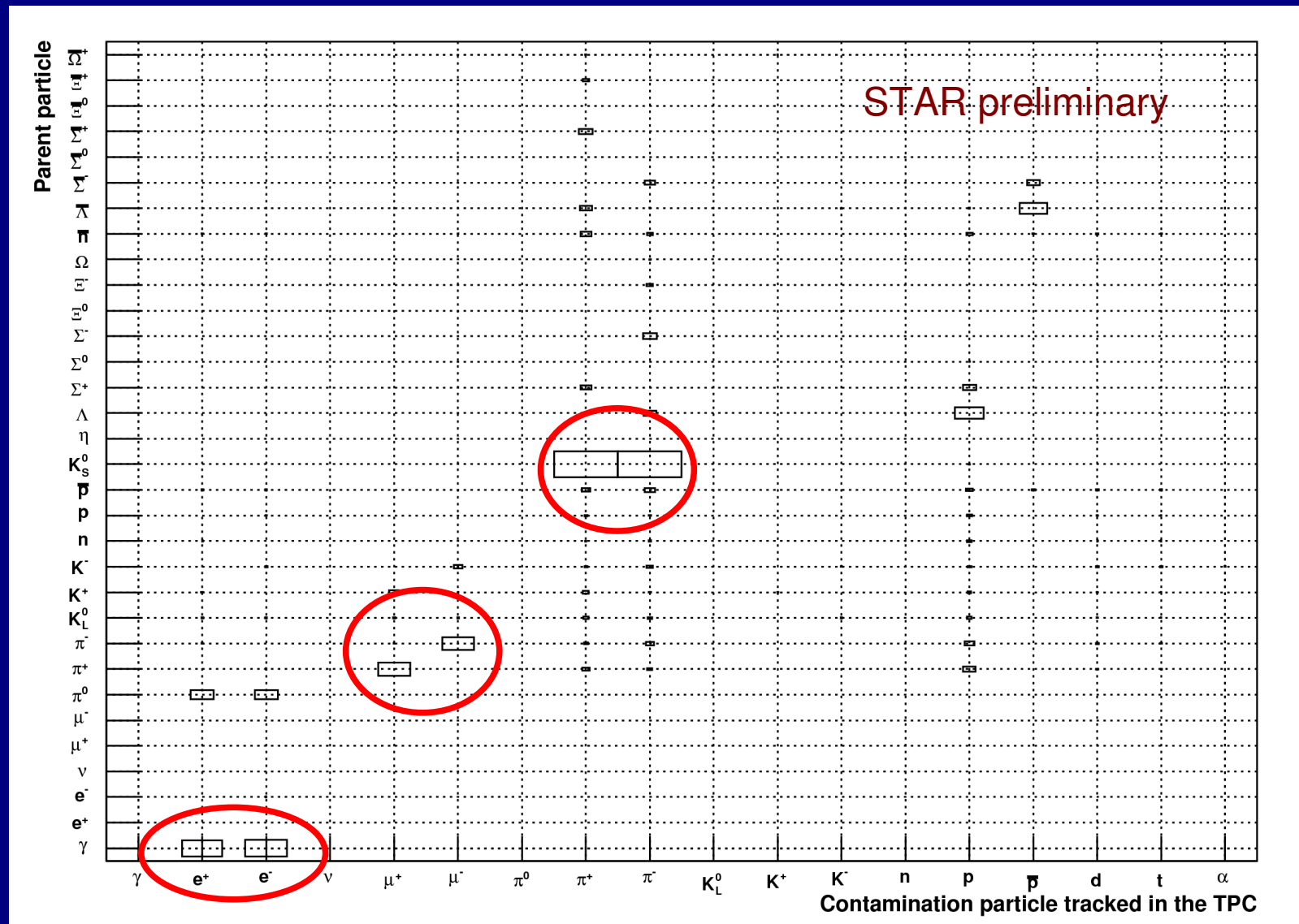
Correcting the measured N_{ch} distribution

- measured distribution: $P(M)$
- corrected distribution: $P(N)$
- the probability $P(M|N)$ obtained from simulation: $P(G|H)$
- need the inverse probability $P(N|M)$, then:

$$P(N) = \sum_M P(N|M) \cdot P(M)$$

- can't simply invert $P(M|N)$ – statistical fluctuations, can be singular
- G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. A **362** (1995) 487.

Bayes' theorem – how to get the inverse probability

unfolding matrix

corrected multiplicity distribution

$$P(N|M) = \frac{P(M|N) \cdot P(N)}{\sum_{i=1}^{n_N} P(M|N_i) \cdot P(N_i)}$$

smearing matrix

iterative approach:

1. start with uniform $P(N)$
2. compute $P(N|M)$ using $P(N)$ and Bayes theorem
3. compute new iteration of $P(N)$ using $P(M)$ and $P(N|M)$

$$P(N) = \sum_M P(N|M) \cdot P(M)$$

4. back to step 2, until it converges ($P(N)$ doesn't change from the previous iteration)

- typically converges after 5-10 iterations
- statistical errors: computed using the final unfolding matrix $P(N|M)$ and measured multiplicity $P(M)$

Systematic uncertainties

- possibly different yields of weakly decaying particles and gammas in (Hijing) simulation compared to the data

- products contribute (if pass the DCA cut) to the "after-GEANT" multiplicity ==> affect the unfolding correction

- connected to DCA distributions & cuts:
  - real primaries: <DCA> = 0.46 cm
  - contamination tracks: <DCA> = 1.1 cm
  - fraction of contamination tracks:
    12 % (DCA < 3 cm), 8% (DCA < 1 cm)

DCA distribution from simulation (real primaries + contamination tracks) & data:
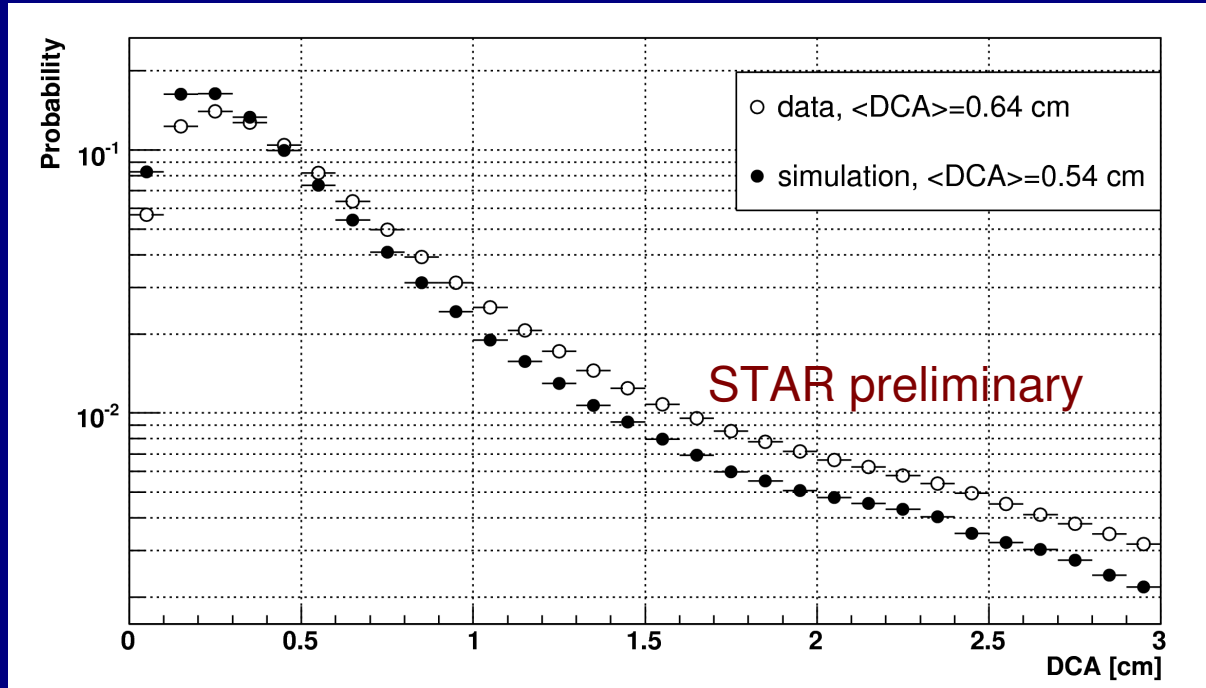
further study needed to match simulation to the data

...so the results depend strongly on DCA cut used:
at the level of (corrected) mean charged multiplicity:

DCA cut [cm]	0.6	1.0	2.0	3.0
$\langle N_{ch} \rangle$	2.23	2.38	2.49	2.54

- explanation: different yields of the particles, whose products cause contamination: mostly K^0_s and gammas
- solution: change (by hand) the yields in the simulation ==> DCA distributions of charged particles are the same between the simulation and the data

Conclusion: systematics not under control yet, we can not show and compare our results now ...

Negative Binomial Distribution (NBD)

$$P(N; \langle N \rangle, k) = \binom{N+k-1}{k-1} \left(\frac{\langle N \rangle / k}{1 + \langle N \rangle / k} \right)^N \frac{1}{(1 + \langle N \rangle / k)^k}$$

broader than Poisson (independent particle production):

$$D = \langle N \rangle + \frac{\langle N \rangle^2}{k}$$

limit cases: k → ∞ Poisson, k → 1 geometric dist.

2 energy-dependent parameters:
<N> increases and *k* decreases with √*s*

Predictions for LHC energies

- from fixed target (11 GeV) through ISR energies to 900 GeV ($\mathrm{S\bar{p}pS}$) N_{ch} distributions follow NBD

- empirical formulas for NBD parameters (UA5):
  - $1/k = \text{alpha} + \text{beta} * \ln(\text{sqrt}(s))$
  - $<N_{ch}> = a + b*\ln(\text{sqrt}(s)) + c*\ln(\text{sqrt}(s))^2$
  - $<N_{ch}> = a + b*(\text{sqrt}(s))^c$

- fitted to UA5 data (200, 546, 900 GeV) for $|\eta|<0.5$, without p_T cut

- $<N_{ch}>$: can't distinguish between power and $\log^2$

Fit results
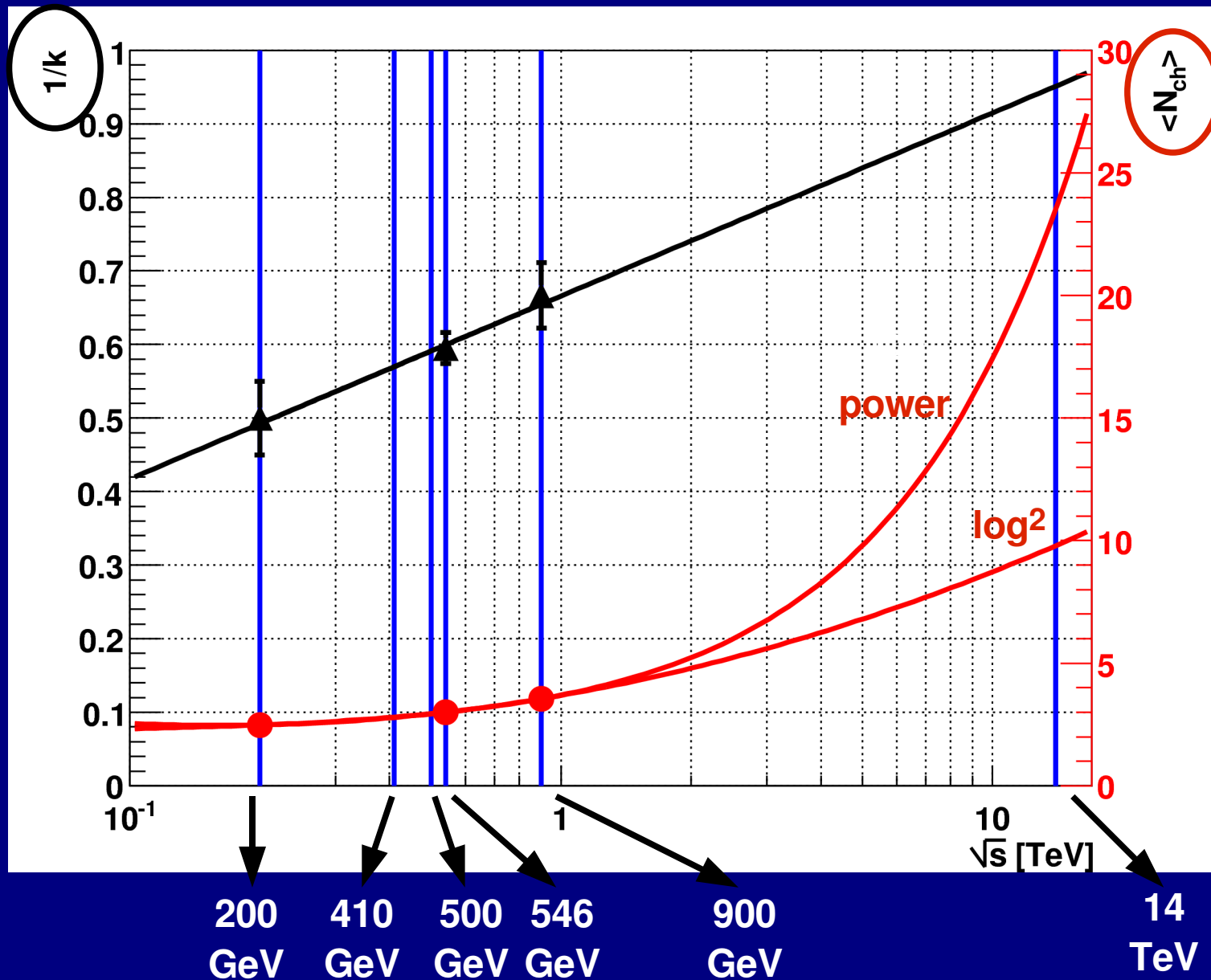
UA5 fitted:
200 GeV
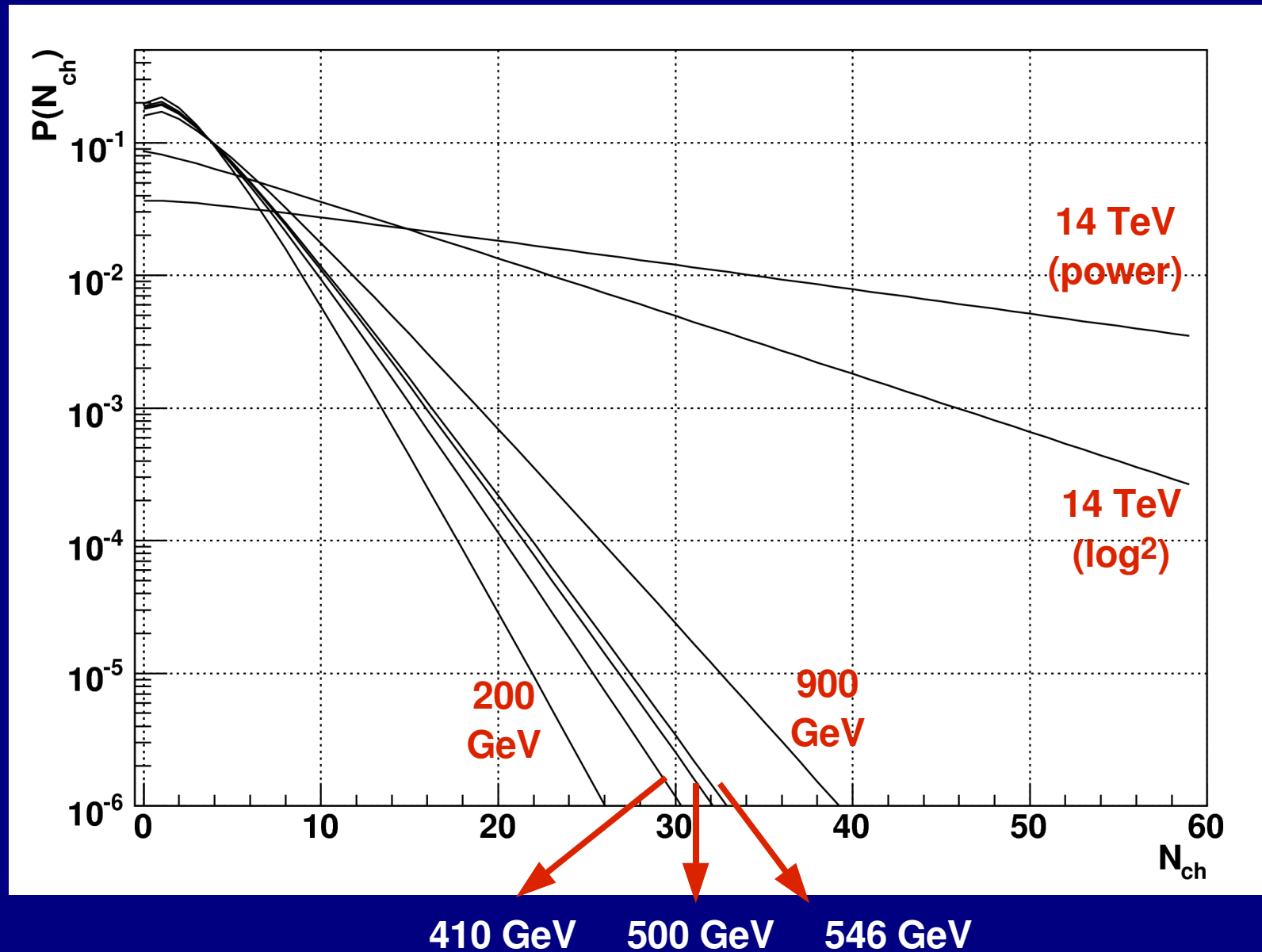546 GeV
900 GeV

RHIC:
200 GeV
410 GeV (test)
500 GeV (future)

LHC:
14 TeV

NBD – predicted N_{ch} distributions

Conclusion

- vertex finding efficiency correction works well
- unfolding correction method works well, but the input from simulation is affected by a large systematic uncertainty (different DCA distributions, connected with different K^0_s and gamma yields)
- except for very high multiplicities (not measured by UA5), we can predict what N_{ch} distributions will look like at LHC

Backup

- z_{vertex} distributions:
- simulation: sigma = 40 cm
- data: sigma = 65 cm
- therefore for high z_{vertex} insufficient statistics in simulation
- for +-25 cm I've got 3.5M in data, 275K in simul.

k and <Nch> from the fits – log;
power-like fit gives 2.80, 2.94 and 23.5

$\sqrt{s}$ [TeV]	$\langle N_{ch} \rangle$	k
0.2	2.48 ± 0.06	2.0 ± 0.2
0.41	2.78	1.76
0.5	2.93	1.69
0.546	3.00 ± 0.04	1.68 ± 0.06
0.9	3.55 ± 0.07	1.5 ± 0.1
14	9.79	1.05

references:

- NBD: G. J. Alner *et al.* [UA5 Collaboration], *A New Empirical Regularity For Multiplicity Distributions In Place Of Kno Scaling*, Phys. Lett. B **160**, 199 (1985).
- UA5 results: R. E. Ansorge *et al.* [UA5 Collaboration], *Charged particle multiplicity distributions at 200 GeV and 900 GeV center-of-mass energy*, Z. Phys. C **43** (1989) 357.
- UA5 results @ 546 GeV: G. J. Alner *et al.* [UA5 Collaboration], *An Investigation Of Multiplicity Distributions In Different Pseudorapidity Intervals In Anti-P P Reactions At A Cms Energy Of 540-GeV,* Phys. Lett. B **160** (1985) 193.*

NBD properties & KNO

KNO scaling --> C_n moments independent of energy

NBD:

parameters N, k

$C_2 = 1 + 1/N + 1/k$

$C_3 = 1 + 3(1/N+1/k) + (1/N+1/k)^2 + 1/k(1/N+1/k)$

etc.

KNO broken due to N and k behaviour versus sqrt(s);
but: if NBD holds, KNO is broken irrespective of N, k behaviour vs.
sqrt(s) – of course, N rises with sqrt(s)
breakdown much more evident at the full phase space (at $|\eta|$<0.5, C_n
moments have big errors)
at sqrt(s) 11-62 GeV (before $S\bar{p}pS$), accidental scaling: 1/N+1/k
approximately independent of sqrt(s), 1/k still quite small at these
energies